Appendix: derivation of sensitivity formula

N=number of women having triple test; p=prevalence of Down's syndrome; $p*N$=total number of fetuses with Down's syndrome; DR=detection rate; FPR=false positive rate; $p*N*DR*AR$=number of women with true positive results who have amniocentesis; $(1-p)*N*FPR*AR$=number of women with false positive results who have amniocentesis; $p*N*DR*AR$=number having an abortion (all positive amniocentesis results); TT=cost of triple test per person; Am=cost of amniocentesis; Ab=cost of abortion; AR=amniocentesis acceptance rate.

The total cost of avoiding the birth of all detected fetuses with Down's syndrome has three components: the cost of the triple test ($TT*N$); the cost of amniocentesis in all women with true positive results or false positive results; and the cost of abortion.

$$\text{Total cost} = TT*N+Am*[p*N*DR+(1-p)*N*FPR]*AR + Ab*p*N*DR*AR$$

$$\text{Average cost} = \text{Total cost/Number of births avoided}$$

$$= \frac{TT*N+Am*[p*N*DR+(1-p)*N*FPR)]*AR + Ab*p*N*DR*AR}{p*N*DR*AR}$$

$$= \frac{TT(1/AR)}{p*DR} + Am + \frac{Am*FPR}{p*DR} + Ab$$

The only term that varies as the acceptance rate (AR) changes is the one on the left. Assuming an acceptance rate of 75% as in table III this term is:

$$\frac{TT(1/AR)}{p*DR} = \frac{TT(1/0{\cdot}75)}{p*DR} = \frac{TT(1{\cdot}33)}{p*DR}$$

Thus changes in acceptance rate affect the average cost according to the formula given in the text:

Change in average cost=(cost per triple test) (1/AR−1·33)/(prevalence rate) (detection rate).

1 Ferguson-Smith M. Neural tube defects and Down's syndrome: a success and a failure. In: *Screening for fetal and genetic abnormality*. London: King's Fund Centre, 1987.
2 Hagard S, Carter FA. Preventing the birth of infants with Down's syndrome: a cost benefit analysis. *BMJ* 1976;i:753-6.
3 Wald NS, Cuckle HS. Recent advances in screening for Down's syndrome and neural tube defects. In: Rodeck C, ed. *Balliere's clinical obstetrics and gynaecology*. Vol 1. No 3. London: Baillière Tindall, 1987: 649-76.
4 McGrother CW, Marshall BS. Recent trends in incidence, mortality and survival in Down's syndrome. *J Ment Defic Res* 1990;**34**:49-57.
5 Wald NJ, Cuckle HS, Densem JW, *et al*. Maternal serum screening for Down's syndrome in early pregnancy. *BMJ* 1988;**297**:883-7.
6 McKeown T. Validation of screening procedures. In: *Screening in medical care: reviewing the evidence, a collection of essays*. London: Oxford University Press, 1968.
7 Henderson JB. Economic evaluation of screening for fetal and genetic abnormality. In: *Screening for fetal and genetic abnormality*. London: King's Fund Centre, 1987.
8 Sadovnic AD, Baird PA. A cost-benefit analysis of prenatal detection of Down's syndrome and neural tube defects in older mothers. *Am J Med Genet* 1981:**10**:367-78.
9 Gill M, Murday V, Slack J. An economic appraisal of screening for Down's syndrome in pregnancy using maternal age and serum alpha fetoprotein concentration. *Soc Sci Med* 1987;**24**:725-31.
10 Callan J. *The economics of pre-natal screening*. York: Centre for Health Economics, University of York, 1988. (Discussion paper No 42.)
11 Drummond MF, Stoddart GL, Torrance GW. *Methods for the economic evaluation of health care programs*. Oxford: Oxford University Press, 1987.
12 Cartmill RS, Thornton JG, Williams J, Holding S, Lilford RJ. Clinical experience with the triple test for Down's syndrome screening. *Journal of Perinatal Medicine* 1991 (in press).
13 Dowie J. *Professional judgement, introductory text* 5. D321. Milton Keynes: Open University, 1987:9.
14 Richards M. Some reflections on fetal diagnosis. In: *Screening for fetal and genetic abnormality*. London: King's Fund Centre, 1987.
15 King's Fund Forum Consensus Statement. *Screening for fetal and genetic abnormality*. London: King Edward's Hospital Fund for London, 1987.
16 Drummond M. Personal view. *BMJ* 1988;**296**:566.
17 Weinstein MC, Fineberg HV, Elstein AS, *et al*. *Clinical decision analysis*. Philadelphia: W B Saunders, 1980.
18 Kassiner JP, Moskowitz AJ, Lau J, Park SG. Decision analysis: a progress report. *Ann Intern Med* 1987;**106**:275-91.
19 Doubilet P, McNeil BJ. Clinical decision making. *Med Care* 1985;**23**:648-62.
20 Schwartz WB. Decision analysis: a look at the chief complaints. *N Engl J Med* 1979;**300**:556-9.
21 Thornton JG. Decision analysis in prenatal diagnosis: measuring patients' values. In: Lilford RJ, ed. *Prenatal diagnosis and prognosis*. London: Butterworths, 1990.
22 Pauker SP, Pauker SG. Prenatal diagnosis: a directive approach to genetic counselling using decision analysis. *Yale J Biol Med* 1977;**50**:275-87.
23 Pauker SP, Pauker SG. The amniocentesis decision: ten years of decision analytic experience. *March of Dimes Birth Defects Foundation, birth defects: original article series*, 1987;**23**:151-69.

(Accepted 21 February 1991)

Critical appraisal of published research: introductory guidelines

F G R Fowkes, P M Fulton

Reading medical journals can be time consuming, and most doctors have become adept at skimming contents by reading titles and perusing abstracts. Getting to grips with the details of papers and conducting a critical appraisal of the research may be a task many find difficult because they are not research experts and have not had training in the appraisal of papers during medical education. This difficulty may be particularly acute for junior doctors embarking on their first research towards a higher degree.

The purpose of appraising a paper is to discover if the methods and results of the research are sufficiently valid to produce useful information. The prime objective is not necessarily to assess the authors; sometimes a research project may be the best which could be carried out but, because of the unforseen difficulties, the results are of limited value. Neither is the purpose to decide if the research is well presented, beautifully written, and finely illustrated. A critical appraisal is concerned with assessing the hard facts of the research.

In this article we set out simple guidelines for appraising medical research on patients or in the general population. These guidelines do not cover all categories of study design, and individual guidelines are not necessarily relevant to every study. Instead they are reminders of the important features which should be considered when appraising a paper. The statistical content of papers has been discussed[1,2] and, although important, is not reviewed here.

Department of Community Medicine, University of Edinburgh, Medical School, Edinburgh EH8 9AG
F G R Fowkes, FRCPED, *reader*
P M Fulton, FRCPED, *senior lecturer*

Correspondence to: Dr Fowkes.

BMJ 1991;**302**:1136-40

Starting the appraisal

REVIEW A STRUCTURED ABSTRACT

An abstract should summarise the research succinctly and include the objectives, study design, important results, and the authors' conclusions. When starting to appraise an article reviewing a structured abstract is particularly helpful because the different components of the research can readily be identified.[3,4] Serious flaws may be detected at this stage.

WHAT ARE THE OBJECTIVES OF THE RESEARCH?

The next step is to identify the precise objectives of the research. These should be stated in the introduction to the paper or may be given in the abstract. A principal goal of appraisal is to establish whether research objectives have been met, and the most precise form of the objectives should be sought. For example, the specific objective "To determine the effect of one week's bed rest, compared with normal activity, on the severity and duration of acute lumbosacral pain" is more useful than the general objective "To examine the role of bed rest in the management of back pain." Unfortunately, specific objectives are often not stated in papers and we may even have to peruse the results in order to find the apparent objectives. Asking ourselves "What are the authors really trying to measure?" can help with this.

WHAT IS THE OVERALL STUDY DESIGN?

Before dissecting the methods in detail the overall design of the study should be clarified, as this helps to determine those aspects of the appraisal on which to concentrate. The design should be stated in the abstract and at the beginning of the methods section. Most studies comprise one of six designs: case report, case series, cross sectional, cohort, case-control, and controlled trial. Many others are simply combinations and nuances of these six. The formats are as follows (figure).

Case report

A case report is a description of one interesting and unusual case.

Case series

A case series is a description of several cases in which no attempt is made to answer specific hypotheses or to compare the results with another group of cases.

Example—Medical records were reviewed to study the clinical features of measles diagnosed in children treated for malignant disease in Newcastle upon Tyne during 1973-86. Of the 17 cases diagnosed, five were fatal. No other cases were studied.[5]

Cross sectional study

A cross sectional study is a survey of the frequency of disease, risk factors, or other characteristics in a defined population at one particular time.

Example—The prevalence of diagnosed and undiagnosed diabetes mellitus was investigated in a predominantly Asian population in Coventry.[6] Residents in one electoral ward were visited at home, where a questionnaire was administered and a capillary blood sample taken. Diabetes occurred more frequently in Asians; undiagnosed disease was found in over a quarter of the population studied.

Cohort study

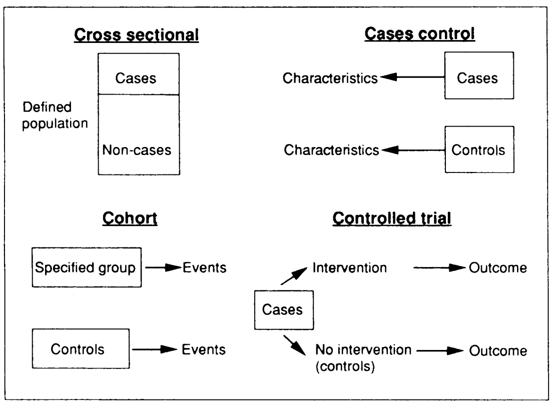
A cohort study is an observational study of a group of people with a specified characteristic or disease who are followed up over a period of time to detect new events. Comparisons may be made with a control group. No interventions are normally applied to the groups.

Example—Two hundred and fifty seven consecutive patients with intermittent claudication were followed up for a mean of 6·5 years. Forty four per cent (113) of the patients died; the mortality was twice that of the general population matched for age and sex.[7]

Case-control study

A case-control study is an observational study in which characteristics of people with a disease (cases) are compared with those of selected people without the disease (controls).

Example—A study was undertaken to test the hypothesis that exposure to dust is a risk factor for cryptogenic fibrosing alveolitis. A questionnaire asking about lifetime exposure to dust was administered to 40 patients (cases) and 106 community controls matched for age and sex.[8]



Formats for four types of study

Controlled trial

A controlled trial is an experimental study in which an intervention is applied to one group of people and the outcome compared with that in a similar group (controls) not receiving the intervention.

Example—A total of 1377 patients with symptomatic atherosclerotic disease of the internal carotid artery were randomly allocated to "best medical care" or the same regimen plus extracranial-intracranial bypass surgery. Mortality and morbidity due to stroke were compared between the two groups during a follow up period of almost five years.[9]

Guidelines

The next step, having identified the study objectives and overall design, is to conduct a detailed appraisal of the methods and results. The following six guidelines, each in the form of a question about the research and including a checklist of criteria, are summarised in the box.

(1) STUDY DESIGN APPROPRIATE TO OBJECTIVES?

Deciding if the overall study design is appropriate may require more common sense than a detailed knowledge of epidemiological methods. If, for example, the purpose of a study is to evaluate a new treatment a controlled trial is almost imperative, as a trial without a control group would be fraught with difficulties in knowing whether improvement in patients was due to the treatment. Similarly, a project examining prognosis would normally require follow up by means of a cohort study. On the other hand, research investigating the cause of disease might adopt any of the designs shown in the figure.

(2) STUDY SAMPLE REPRESENTATIVE?

Source of sample

If research is to be applicable and relevant to other populations the study sample (group selected to participate) must be representative of the group from which it is drawn (study population), which in turn should be typical of the wider population to whom the research might apply (target population). Appropriateness of the target and study populations is usually a subjective assessment based on our knowledge of the topic under investigation. For example, research concerned with the pathogenesis of coronary artery disease might be of limited value if restricted to a target population of women over 80 years of age.

Sampling method

In population based studies random sampling is the ideal method of avoiding selection bias and producing a sample typical of the study population. In other studies non-random sampling may be adequate; for example, consecutive patients attending a clinic may be included in a controlled trial, or every nth person may be selected from a register. In studies based in hospital, however, beware that referral bias may lead to an atypical study sample.

Sample size

A statement in the methods section that a sample size was chosen in order to have sufficient power to detect a medically meaningful result at a certain level of statistical significance would normally be adequate evidence that steps had been taken to ensure an appropriate sample size. In the absence of such a statement it may be necessary to seek help from a statistician or an appropriate text[10,11] to establish whether the sample size was adequate. But it is also important to assess the biological representativeness of the sample. Was the sample large enough to encompass the full range of disease? Or was it so small that there was a danger of a biased homogeneous group having a disproportionate effect on the results? It is not uncommon, for example, to read of statistically valid randomised controlled trials containing fewer than 20 patients. Was it likely with such a small number of patients that they were truly representative of all those presenting to clinicians in other centres?

Entry criteria and exclusions

The criteria for entering subjects into a study must be examined carefully; the stage of disease or time of onset, for example, may have a profound effect on the results of treatment or in the detection of aetiological factors. Exclusion criteria should also be defined appropriately. Furthermore, any description of the study participants must be scrutinised in order to assess whether the sample was representative.

Non-respondents

In most studies some subjects do not respond to invitations, some refuse to participate, and others do not attend for examination. The response rate is often viewed as an indicator of the representativeness of participants, but the size of response is only one aspect of sampling and may be less important than the comparability between participants and non-respondents. For example, a response rate of 30% may be satisfactory if there is good evidence that participants do not have atypical characteristics which might affect the results of the research. Thus comparisons should be sought between participants and the non-respondents or the total study population.

EXAMPLE

Collin *et al* carried out a community screening programme for abdominal aortic aneurysm in men aged 65-74.[12] Invitations to attend hospital for an examination were sent to 843 men identified from the age-sex registers of two large group practices. Four hundred and twenty six men attended, giving a response rate of 50·5%. Although the authors' main interest was in those who attended, the respondents may have differed from the non-respondents in important respects—for example, past medical history, current illness, and socioeconomic state. Thus scope for extrapolating the results to a wider population was limited.

(3) CONTROL GROUP ACCEPTABLE?

Definition of controls

In studies using a comparison or control group it is important to assess whether this group was adequate for the purpose under study. In a case-control study, for example, were the criteria for defining controls appropriate and was the control group checked to ensure that it did not contain cases?

Source of controls

In case-control and cohort studies the source of controls should be such that the distributions of characteristics (not under direct investigation) are similar to those in the cases or study cohort. For example, in a study of exposure to lead and mental ability in children the source of controls should ideally be a group whose social class distribution is similar to that from which the cases were derived.

Matching and randomisation

In case-control studies cases and controls are often matched for certain characteristics, such as age and sex. Did the matching process seem to have been carried out correctly? In controlled trials, on the other hand, subjects are often randomly allocated to intervention and control groups. The method of randomisation should be assessed to ensure that the subjects were truly randomised—for example, by use of computer generated random numbers.

Comparable characteristics

In controlled trials random allocation to intervention and control groups usually leads to comparability, but not necessarily so, and the distributions of age, sex, and other prognostic variables should therefore be compared between the two groups. Similarly, in case-control and cohort studies matching or other methods of selecting controls are not infallible and the comparability of the groups must be assessed.

EXAMPLE

In a case-control study in Adelaide of diet, alcohol, and weight in relation to gall stone disease 267 cases were compared with 359 hospital controls and 241 community controls, who were selected at random from the electoral register.[13] The authors pointed out that the diet and lifestyle of the hospital controls were probably atypical because many had gastrointestinal disease. Also, the community controls were probably of a higher socioeconomic state because the cases were selected only in public hospitals whereas some of the controls participated in private health care schemes. But these socioeconomic differences had no effect on the results of the study.

(4) QUALITY OF MEASUREMENTS AND OUTCOMES?

Validity

It is important to assess the validity of measurements made in a research study—that is, the extent to which they reflect the true situation. Dietary questionnaires, for example, are notoriously inaccurate in obtaining a true picture of a person's regular nutritional intake. When a single test is used as a proxy measure of disease the validity of the test (sensitivity and specificity) should be stated in the article. In a randomised controlled trial the results may depend on the measurement of one outcome and it is thus essential that this is an important end point which is sensitive to change.

Reproducibility

In the interests of expediency many research projects pay too little attention to the reproducibility of the measurements. Would the same results have been obtained if the measurements had been taken by a different observer or on a different day? In many larger projects repeatability checks are made at intervals to assess the consistency of measurement. For example, split blood samples may be sent to the laboratory without an indication that they are from the same subject. Evidence on the repeatability of the principal measurements should be sought in the article.

Blindness

During data collection a common source of bias is that the subjects or those collecting the data are not blind to the purpose of the research. The problems that may occur in controlled trials are well known: subjects, observers, and researchers, by wishing the intervention to succeed, produce an unrealistically good success rate. Inadequate blindness may be a problem in other studies. In case-control studies, for example, patients (cases) who are aware of a possible relation between a risk factor and the disease may overreport the risk factor in themselves. Similarly, an observer may make greater efforts to detect a possible risk factor in cases than in controls, or may even unconsciously slant the questions in questionnaires to obtain the desired response. Clearly, in many studies total blindness is

Guidelines and checklist for appraising a medical article

Guideline	*Checklist*	
(1) Study design appropriate to objectives?	Objective:	Common design:
	Prevalence	Cross sectional
	Prognosis	Cohort
	Treatment	Controlled trial
	Cause	Cohort, case-control, cross sectional
(2) Study sample representative?	Source of sample	
	Sampling method	
	Sample size	
	Entry criteria/exclusions	
	Non-respondents	
(3) Control group acceptable?	Definition of controls	
	Source of controls	
	Matching/randomisation	
	Comparable characteristics	
(4) Quality of measurements and outcomes?	Validity	
	Reproducibility	
	Blindness	
	Quality control	
(5) Completeness?	Compliance	
	Drop outs	
	Deaths	
	Missing data	
(6) Distorting influences?	Extraneous treatments	
	Contamination	
	Changes over time	
	Confounding factors	
	Distortion reduced by analysis	

+ + =Major problem. + =Minor problem. 0=No problem. NA=Not applicable.

not feasible, but for the purposes of appraisal it is necessary to consider how this might put the results in doubt.

Quality control

Overall, the extent to which the researchers have instituted quality control measures for the examination of subjects, collection of data, and laboratory tests should give some idea of the likely quality of data. Measures might include testing the accuracy and repeatability of observers, checking the calibration and accuracy of instruments, and random checks for errors in data recording. Laboratories often participate in external quality control schemes, but many clinical researchers do not give adequate attention to this concept.

EXAMPLE

In a retrospective survey information on the symptoms, signs, clinical investigations, and outcomes of 1442 patients with mild head injury admitted to a neurosurgical unit were abstracted from medical records.[14] Although the quality of data may have been satisfactory, there may have been deficiencies in the completeness and accuracy of the medical records and observer bias in detecting abnormalities in the records of patients with poor outcomes. Studies in which data are abstracted from medical records are very prone to such errors.

(5) COMPLETENESS?

Compliance

The end results of a study may be incomplete in relation to the number of subjects who were first enrolled. This need not necessarily lead to bias in the results, but careful assessment is required. In controlled trials continuing compliance of subjects with a regimen may be a serious problem and, although this may partly be overcome by carrying out an "intention to treat" analysis (in which the outcomes of all subjects entering the trial are included in the analysis irrespective of compliance with treatment), when appraising the study it may be quite difficult to assess whether the treatment worked.

Drop outs and deaths

In cohort studies as well as in controlled trials drop outs and deaths in the study sample may occur. It is important to assess not only the proportion of drop outs in each group but also why they dropped out, as this may give a clue to possible bias. For example, more healthy people may move and be lost to follow up, so that a cohort study excluding them might produce an unrealistically gloomy outcome.

Missing data

Incomplete results may often occur due to difficulties in obtaining specimens, laboratory tests going awry, and lost data. The extent and nature of the loss must be assessed in order to estimate possible bias. Also, selectivity in reporting of results and the exclusion of data from tables may have an effect on the conclusions that can be drawn from the research. It is worth checking that in addressing the objectives of the study the authors have presented data on the most appropriate measurements and that some have not mysteriously disappeared.

EXAMPLE

In a cohort study of 5362 subjects born in one week in 1946 blood pressure was measured at 36 years of age to determine associations with social and family factors, smoking, and body mass.[15] A blood pressure measurement was obtained in only 3322 subjects (62%). This substantial loss could have biased the results, but it was shown in comparisons with other data that the cohort was still representative of native born men and women of that generation.

(6) DISTORTING INFLUENCES?

Extraneous treatments

The results of studies are often distorted by outside influences. In controlled trials, for example, a common problem is that subjects may be exposed to treatments in addition to the one being evaluated. Thus in assessing a trial the question has to be asked, "Could there possibly be extraneous treatments which might have influenced the results? Have these been identified in the study and the results interpreted accordingly?"

Contamination

Another problem in controlled trials is contamination, in which one group is affected by another. For example, in a dietary intervention study people in a control group may change their diet because they hear about supposed benefits from dietary changes in the intervention group.

Changes over time

Be wary of studies in which data on a characteristic have been collected from two groups of subjects at different times. Observed differences between the groups might be due to changes in the characteristic or its measurement over time, and not to real differences between the groups.

Confounding factors

Distorting influences may exist in studies examining the association between a risk factor and disease where the purpose is to find out whether the association is real or spurious (caused by a confounding factor influencing both the risk factor and the disease). In such studies it is necessary to account for possible confounding factors. This may be satisfied by matching in the selection of controls or by evidence of comparability between cases and controls.

Distortion reduced by analysis

Distorting influences may also be minimised by some form of stratification or adjustment procedure in the analysis. For example, if smoking is believed to be a confounding factor the results can be examined separately in smokers and non-smokers (stratification) or the results can be adjusted by calculations which take account of different smoking habits (standardisation).[16] Age and sex are frequent confounding factors and invariably should be accounted for by describing age standardised, sex specific rates. Multiple regression is a statistical technique which is often used to analyse independent associations of variables while taking account of confounding factors.[16] In controlled trials outcome measures may have to be analysed separately within subgroups—for example, those exposed and not exposed to extraneous treatments.

EXAMPLE

In a randomised controlled trial a high fibre diet and certain minor surgical procedures were compared in the treatment of haemorrhoids.[17] Contamination may have occurred because patients in the surgical groups could have changed their diet. Also, an unknown number of subjects may have had extraneous treatments, such as topical ointments, sitz baths, and stool softeners. Information was not collected on these possible sources of bias, so that the authors were not able to make adjustments in the analysis and interpretation of the results was difficult.

Making a judgment

Once a detailed appraisal of the methods and results has been conducted a decision must be made on whether the methods were adequate and the results clear cut enough so that the objectives were achieved and useful information produced. Unfortunately, there is no magical formula which will convert assessments of detail into an overall score on the worth of a paper. The pros and cons of the research have to be weighed implicitly and a judgment made. This is one reason why there is such scope for diametrically opposed views to be expressed in the correspondence columns of journals.

Some aspects of study design may have a more important influence than others but it is impossible to be categorical as much depends on the objectives and overall study design. For example, in a trial deficiencies in the allocation of controls would probably be more important than inadequate evidence on the reproducibility of measurements. When checking the criteria for each guideline, as shown in the box, assigning problems for each criterion as major (++) or minor (+) in terms of their expected effect on the results may be helpful in drawing conclusions.

In attempting to sum up a paper it may be helpful to ask three questions:

(1) *Bias*—Are the results erroneously biased in a certain direction? This may not necessarily negate the value of a study as long as the direction and magnitude of the bias are known.

(2) *Confounding*—Are there any serious confounding or other distorting influences? Often these cannot be adequately accounted for in the analysis and may have a substantial effect on the results.

(3) *Chance*—Is it likely that the results occurred by chance? The answer depends primarily on appraisal of the statistical content,[1,2] and help from a statistician may be required.

If the answer to each question is categorically "No" the research is probably quite sound.

In conclusion, conducting a critical appraisal of a paper is a worthwhile task but the overall judgment is often difficult. Papers are rarely judged to be "very good" or "very bad" but usually lie on a continuum in between. Most are likely to be of some value but accompanied by reservations—"This study has produced some interesting results but has its problems."

We are grateful to the following for comments on this paper: Mr P T Donnan, Dr P Flanagan, Professor W M Garraway, Dr E Housley, Ms P Isard, Dr G C Leng, Mr C V Ruckley, Ms F B Smith, Professor C P Warlow, Mr M R Whyman.

1 Altman DG, Gore SM, Gardner MJ, Pocock SJ. Statistical guidelines for contributors to medical journals. *BMJ* 1983;**286**:1489-93.
2 Gardner MJ, Machin D, Campbell MJ. Use of check lists in assessing the statistical content of medical studies. *BMJ* 1986;**292**:810-2.
3 Huth EJ. Structured abstracts for papers reporting clinical trials. *Ann Intern Med* 1987;**106**:626-7.
4 Lock S. Structured abstracts. *BMJ* 1988;**297**:156.
5 Kernahan J, McQuillin J, Craft AW. Measles in children who have malignant disease. *BMJ* 1987;**295**:15-8.
6 Simmons D, Williams DRR, Powell MJ. Prevalence of diabetes in a predominantly Asian community: preliminary results of the Coventry diabetes study. *BMJ* 1989;**293**:18-21.
7 Jelnes R, Gaardsting O, Jensen KH, Baekgaard N, Tonnesen KH, Schroeder T. Fate in intermittent claudication: outcome and risk factors. *BMJ* 1986;**293**:1137-40.
8 Scott J, Johnston I, Britton J. What causes cryptogenic fibrosing alveolitis? A case-control study of environmental exposure to dust. *BMJ* 1990;**301**:1015-7.
9 Extracranial/Intracranial Bypass Study Group. Failure of extracranial-intracranial arterial bypass to reduce the risk of ischemic stroke. *N Engl J Med* 1985;**313**:1191-200.
10 Altman DG. Statistics and ethics in medical research. III. How large a sample? *BMJ* 1980;**281**:1336-8.
11 Daly LE. Confidence intervals and sample sizes: don't throw out all your old sample size tables. *BMJ* 1991;**302**:333-6.
12 Collin J, Araujo L, Walton J, Lindsell D. Oxford screening programme for abdominal aortic aneurysm in men aged 65 to 74 years. *Lancet* 1988;ii:613-5.
13 Scragg RKR, McMichael AJ, Baghurst PA. Diet, alcohol, and relative weight in gall stone disease: a case-control study. *BMJ* 1984;**288**:1113-9.
14 Mendelow AR, Campbell DA, Jeffrey RR, *et al.* Admission after mild head injury: benefits and costs. *BMJ* 1982;**285**:1530-2.
15 Wadsworth MEJ, Cripps HA, Midwinter RE, Colley JRT. Blood pressure in a national birth cohort at the age of 36 related to social and familial factors, smoking, and body mass. *BMJ* 1985;**291**:1534-8.
16 Hennekens CH, Buring JE. *Epidemiology in medicine.* Boston: Little, Brown, 1987.
17 Keighley MRB, Buchmann P, Minervini S, Arabi Y, Alexander-Williams J. Prospective trials of minor surgical procedures and high fibre diet for haemorrhoids. *BMJ* 1979;ii:967-9.

FURTHER READING

Cuddy PG, Elenbaas RM, Elenbaas JK. Evaluating the medical literature. *Ann Emerg Med* 1983;**12**:549-55, 610-20, 679-86.
Department of Clinical Epidemiology and Biostatistics, McMaster University Health Sciences Centre. How to read clinical journals. *Can Med Assoc J* 1981;**124**:555-8, 703-10, 869-72, 985-90, 1156-62.
Gehlbach SH. *Interpreting the medical literature. Practical epidemiology for clinicians.* 2nd ed. New York: Collier MacMillan, 1988.
Haynes RB, McKibbon KA, Fitzgerald D, Guyatt GH, Walker CJ, Sackett DL. How to keep up with the medical literature. *Ann Intern Med* 1986;**105**:149-53, 309-12, 474-8, 636-40, 810-6, 978-84.
Last JM, ed. *A dictionary of epidemiology.* 2nd ed. Oxford: Oxford University Press, 1988.
Riegelman RK. *Studying a study and testing a test: how to read the medical literature.* 2nd ed. Boston: Little Brown, 1989.

(*Accepted 1 March 1991*)

ANY QUESTIONS

A woman in her 30s had a patent ductus arteriosus ligated at the age of 16 and has had conflicting advice about the need for antibiotic prophylaxis for dental work. Is it wise to take a prophylactic antibiotic?

Ligation of a patent ductus arteriosus can be regarded as almost the only truly curative cardiac operation. If the duct has been completely obliterated so that there are no longer any abnormal flows or turbulence there is no need for antibiotic prophylaxis. The risk of infective endocarditis developing after bacteraemia is determined by the presence of high velocity jets across pressure drops in the circulation. This was argued long ago by Rodbard and has stood the test of time.[1] Thus, uncorrected secundum atrial septal defect confers a very low risk, small ventricular septal defect a high risk, big ventricular septal defect a low risk, and so on.—CELIA OAKLEY, *consultant cardiologist, London*

1 Rodbard S. Blood velocity and endocarditis. *Circulation* 1963;**27**:18.